

第五章 检验方法的分类及标度

第一节 感官检验方法的选择及应用

一、感官检验的定义及目的

所谓食品感官检验，就是以心理学、生理学、统计学为基础，依靠人的感觉（视、听、触、味、嗅觉）对食品进行评价、测定或检验并进行统计分析以评定食品质量的方法。

对于食品而言，只注重其的营养价值还远远满足不了人们的需要。加工的食品是否味美、人们是否喜欢吃，即加工食品是否满足人们的嗜好，是评定其质量的重要因素之一。对食品成分进行分析，通过测定其中的蛋白质、脂肪、食盐、糖等含量，能够计算其营养价值，但这些并不能说明人们对该食品的嗜好程度，即使是测定食品的粘性、弹性、硬度、酥脆性等物性参数，也不一定能得到和嗜好程度完全一致的数据。因此，即使在物理化学等测试技术和手段飞速发展的今天，对味香及嗜好度等本质上为主观特性的测定，也不得不依靠人的感官检验。人类对食品的最终评价，不借助人而只靠仪器评价，就是在将来也是有一定困难的。可见，食品的感官评价对于评定食品的质量是十分重要，也是十分必要的。感官评价除了可以评定食品的质量外，在食品生产过程中，还可以利用感官检验方法从食品制造工艺的原材料或中间产品的感官特性来预测产品的质量，为加工工艺的合理选择，正确操作，优化控制提供有关的数据，以控制和预测产品的质量和顾客对产品的满意程度。因此，感官检验对产品质量的预测和控制也具有重要作用。

然而，由于人们对食品的嗜好千差万别，即使是同一个人，也因其心理状态、生理状态及环境的变化，对同一种食品的嗜好表现通常也是不一样的。因此，即使是专家所评定的结果，也不一定能代表大多数人的嗜好。食品感官检验主要是研究怎样从大多数食用者当中选择必要的人选（称评判员），在一定的条件下对试样加以品评，并将结果填写在问答票（评分单）中，然后对他们的回答结果进行统计分析来客观地评定食品的质量。可见，食品的感官检验绝不是简单的品尝，对于试样、评判员、环境等很多方面均具有严格的规定，根据测试目的和要求的不同，要采用不同的感官检验方法加以实施。

二、感官检验方法分类及其应用

1、感官检验的方法

食品感官检验的方法分为分析型感官检验和嗜好型感官检验两种：

分析型感官检验——把人的感觉作为测定仪器，测定食品的特性或差别的方法。比如：检验酒的杂味，判断用多少人造肉代替香肠中的动物肉人们才能识别出它们之间的差别，评定各种食品的外观、香味食感等特性都属于分析型感官检验。

嗜好型感官检验——根据消费者的嗜好程度评定食品特性的方法。比如：饮料的甜度怎样算最好，电冰箱颜色怎样最好等。

弄清感官检验的目的，分清是利用人的感觉测定物质的特性（分析型）还是通过物质来测定人们嗜好度（嗜好型）是设计感官检验的出发点。例如，对二种冰淇淋，如果要研究二者的差别，就可以把冰淇淋溶解或用水稀释，应在最容易检查出其差别的条件下进行检验，但如果要研究哪种冰淇淋受消费者欢迎，通常必须在一般能吃的状态下进行检验。

2、常用试验方法

根据感官检验工作的目的和要求，常用的试验方法有以下六种：

（1）差别试验（Difference Test）

差别试验用于分辨样品之间的差别，其中包括 2 个样品或者是多个样品之间的差别试

验。

差别试验是对样品进行选择性的比较,一般领先于其它试验,在许多方面有广泛的用途。例如在贮藏试验中,可以比较不同的贮藏时间对食品的味觉、口感、鲜度等质量指标的影响。又如在外包装试验中,可以判断哪种包装形式更受欢迎,而成本高的包装形式有时并不一定受消费者欢迎,都可以用差别试验检验。

对于 2 个样品的比较,可以直接判断出它们之间是否存在着差别。其试验方法为 2 点比较法,试验形式有 AB、BA、AA、BB 组合,每次试验中,每个样品的猜测性(有无差别)的概率为 $1/2$ 。如果增加试验次数至 n 次,那么其概率将降低至 $1/2^n$ 。所以在条件许可的情况下,应尽可能增加试验次数。有时在试验中也可以以某个样品(已知)作为标准品,再对其它样品进行配对比较。从而判断出它们之间有无差别及差别程度,其试验方法有 AAB、ABA、ABB、BAA、BBA、BAB 等 6 种。每次试验中,每个样品猜测性的概率值为 $1/3$ 。试验次数的增加会降低其猜测性。当猜测性的概率值小于 5% 时,试验次数分别应不小于 5 次(2 个样品之间)和 3 次(3 个样品之间),见表 5-1。

表 5-1 试验次数对猜测性的影响

猜测概率	试验次数 (n)					
	1	2	3	4	5	6
$1/2^n$	0.5	0.25	0.13	0.063	0.031	0.016
$1/3^n$	0.33	0.11	0.036	0.012	0.0039	0.0013

差别试验的试验方法有:2 点识别法、2 点嗜好法、1:2 点比较法、3 点识别法、3 点嗜好法等。试验结果的分析常用查表法。

①2 点比较法是比较两种试样,以此来区别两者或判断其优劣的方法。这是最简单,最基本的方法。可按试验目的分为 2 点识别法和 2 点嗜好法。

2 点识别法是比较 X、Y 两种试样,根据人的感觉排列 X、Y 的顺序,即区别两者的方法。由于 X 和 Y 之间的顺序是客观存在的。当人们的感觉判断的顺序和客观存在的顺序一致时,回答是正确的,否则回答是错误的,因此识别检验只需做单侧检验。2 点识别法一般用于判断评审员的识别能力或者判断 X、Y 之间的差别是否达到能识别的程度等。本法具有准备和实施方便等优点,缺点是结果差错的偶然可能性大。

2 点嗜好法是指比较 X、Y 两种试样后指出自己喜欢哪一种的方法。在嗜好性检验中,评审员指出 X、Y 两种试样中的任何一个均可以,故必须进行双侧检验。本法主要用于市场调查和质量检验。

②1:2 比较法

1:2 比较法是指先供给试样 X,让评审员记住它的特性(这个试样称明试样),然后同时供给用暗号表示的试样 X 和 Y(因为评审员事先不知两个试样的内容和特性,故称暗试样),让评审员判断两个暗试样中哪个是试样 X 的试验。1:2 比较法一般用于出厂检查验收商品,或用于测定评审员的识别能力,该法比 2 点比较法灵敏度高。

③3 点比较法

有 2 个试样 X、Y,把两个相同的试样和一个不同的试样按 XYY、XXY、XYX 等方式组合后供给评判员,让评审员判断其中一个不同的试样的方法叫做 3 点识别法。然后再比较一个试样和剩余的两个相同试样,判断喜欢哪一个的方法叫做 3 点嗜好法。因此,3 点比较法只

经一次试验，就能同时完成识别和嗜好两个试验。

（2）排列试验（Ranking Test）

排列试验对某种食品的质量指标，按大小或强弱顺序对样品进行排列，并记上 1、2、3……数字。它具有简单并且能够评判 2 个以上样品的特点。其缺点是排列试验只是一个初步的分辨试验形式，它无法判断样品之间差别大小和程度，只是其试验数据之间进行比较。试验结果的分析常用查表法和方差分析法。

（3）分级试验（Scoring Test）

分级试验按照特定的分级尺度，对试样进行评判，并给以适当的级数值。

分级试验是以某个级数值来描述食品的属性。在排列试验中，两个样品之间必须存在先后顺序，而在分级试验中，两个样品可能属于同一级数，也可能属于不同级数，而且它们之间的级数差别可大可小。排列试验和分级试验各有特点和针对性。

分级试验的试验方法主要有评分法、scheffe 一对比较法、模糊数学法等。试验结果的分析常用方差分析法。

（4）阈值试验（Threshold Test）

阈值试验是通过稀释（样品）确定感官分辨某一质量指标的最小值。

阈值试验主要用于味觉的测定，测定值有：

①刺激阈（RL）

能够分辨出感觉的最小刺激量叫做刺激阈。刺激阈分为：敏感阈、识别阈和极限阈。阈值大小取决于刺激的性质和评价员的敏感度，阈值大小也因测定方法的不同而发生变化。

②分辨阈（DL）

感觉上能够分辨出刺激量的最小变化量称分辨阈。用 $\pm \Delta S$ 来表示刺激量的增加(上)或减少(下), 上下分辨阈的绝对值的平均值称平均分辨阈。

③主观等价值（DSE）

对某些感官特性而言，有时两个刺激产生相同的感受效果，我们称之为等价刺激。例如：10%的葡萄糖与 6.3%的蔗糖的刺激等价。

阈值试验的试验方法主要有极限法和定常法。

关于阈值试验更加详细的内容，我们将在第八章中加以详细介绍。

（5）分析或描述试验（Analysis or Description Test）

描述试验是对样品与标准样品之间进行比较，给出较为准确的描述。

描述试验要求试验人员对食品的质量指标用合理、清楚的文字作准确的描述。描述试验有颜色和外表描述、风味描述、质构描述和定量描述。其主要用途有：新产品的研制与开发；鉴别产品间的差别；质量控制；为仪器检验提供感官数据；提供产品特性的永久记录；监测产品在贮藏期间的变化等。

因为感官感觉中任何一个器官的机能活动，不仅取决于直接刺激该器官所引起的响应，而且还受到其它感觉系统的影响，即感觉器官之间相互联系、相互作用。所以，食品的感官感觉是不同强度的各种感觉的总和。并且各种不同刺激物的影响性质各不相同，因此，在食品感官检验中，即要控制一定条件来恒定一些因素的影响，又要考虑各种因素之间的互相关联作用。目前常用的分析和描述性检验方法主要有：简单描述检验法及定量描述和感官剖面检验法。

①简单描述检验法

它是评价员对构成产品特性的各个指标进行定性描述，尽量完整地描述出样品品质的检验方法，描述检验按评价内容可分为风味描述和质地描述。

按评价方式可分为自由式评价和界定式描述。自由式描述即评价员可用任意的词汇，对样品特性进行描述，但评价员一般需要对产品特性非常熟悉或受过专门训练；界定式描述则

在评价前由评价组织者提供指标检验表，评价员是在指标检验表的指导下进行评价的。

该方法多用在食品加工中质量控制，产品贮藏期间质量变化，以及鉴评员培训等情况。

最后，在完成鉴评工作后，要由评价小组组织者统计结果，并将结果公布，由小组讨论确定鉴评结果。

②定量描述和感官剖面检验法

它是评价员尽量完整地描述食品感官特性以及这些特性强度的检验方法。

这种方法多用于产品质量控制、质量分析、判定产品差异性、新产品开发和产品品质改良等方面，还可以为仪器检验结果提供可对比的感官数据，使产品特性可以相对稳定地保存下来。

这种方法依照检验方式的不同可分为一致方法和独立方法二大类型。一致方法的含义是，在检验中所有的评价员（包括评价小组组长）以一个集体的一部分而工作，目的是获得一个评价小组赞同的综合印象，使描述产品风味特点达到一致、获得同感的方法。在检验过程中，如果不能一次达成共识，可借助参比样来进行，有时需要多次讨论方可达到目的。独立方法是由评价员先在小组内讨论产品的风味，然后由每个评价员单独工作，记录对食品感觉的评价成绩，最后用计算的平均值的方法，获得评价结果。无论是一致方法还是独立方法，在检验开始前，评价组织者和评价员应完成以下工作：①制定记录样品的特殊目录；②确定参比样；③规定描述特性的词汇；④建立描述和检验样品的方法。该方法的具体介绍将放在第九章。

（6）消费者试验（Consumer Test）

消费者试验是由顾客根据各人的爱好对食品进行评判。

生产食品的最终目的是使食品被消费者接受和喜爱。消费者试验的目的是确定广大消费者对食品的态度。主要用于市场调查、向社会介绍新产品、进行预测等。

由于消费者一般都没有经过正规培训，个人的爱好、偏食习惯、感官敏感性等情况都不一致，故要求试验形式尽可能简单、明了、易行。使得广大消费者乐于接受，而且要保证参加人数较多（50~80）。

3、试验方法的应用选择及问答票

（1）感官试验方法的应用选择

根据感官检验工作的目的和要求，可以选择适当的试验方法。试验方法的选择主要取决于食品的性质和评审员两方面的因素。例如，对于辣味和刺激性比较强的食品，应该选择差别试验方法比较合适，这样可以避免因为多次品尝而引起的感觉疲劳。表 5-2 列出了感官检验方法的选择，表 5-3 列出了不同检验方法所需评价人数。

表 5-2 感官检验方法的选择

实际应用	检验目的	方法
生产过程中的质量控制	检出与标准品有无差异	成对比较检验法(单边) 成对比较检验法(双边) 二-三点检验法 三点检验法 选择法 配偶法
	检出与标准差异的量	评分法 成对比较检验法 三点检验法

原料品质控制检查	原料的分等	评分法 分等法（总体的）
成品质量控制检查	检出趋向性差异	评分法 分等法
消费者嗜好调查，成品品质研究	获知嗜好程度 或品质好坏	成对比较检验法 三点检验法 排序法 选择法
	嗜好程度 或感官品质顺序 评分法的数量化	评分法 多重比较法 配偶法
品质研究	分析品质内容	描述法

表 5-3 不同检验方法所需评价人数

方 法	所需评价员人数		
	专家型	优秀选评员	初级评价员
成对比较检验法	7 名以上	20 名以上	30 名以上
三点检验法	6 名以上	15 名以上	25 名以上
二-三点检验法			20 名以上
五中取二检验法		10 名以上	
“A”或非“A”检验法		20 名以上	30 名以上
排序检验法	2 名以上	5 名以上	10 名以上
分类检验法	3 名以上	3 名以上	
评估检验法	1 名以上	5 名以上	20 名以上
评分检验法	1 名以上	5 名以上	20 名以上
分等检验法	按具体分等方法定	按具体分等方法定	
简单描述检验法	5 名以上	5 名以上	
定量描述或感官剖面检验法	5 名以上	5 名以上	

（2）问答票

在试验过程中，向评审员提出什么样的提问是决定感官检验研究价值的出发点。对于不完备的、无用的提问，无论怎样进行分析，所得的数据结果是毫无意义的。所以在认真品尝和检验样品的基础上，科学地设计问答票是非常重要的。问答票中的问题要明确，避免难于理解和同时有几种答案的提问，提问不应有理论上的矛盾，不应产生诱迫答案，提问不要太多。最好是在正式试验之前先选择几个人进行预备试验，征求他们对问答票的意见。经过综合分析后，再确定问答票的形式及内容。

三、感官检验的常用术语

在食品感官鉴评中，食品的各项感官特性是通过语言表述出来的，而语言本身受到本民族的历史和地域文化的影响，因此很难准确地把握不同国家和不同地区的词语含义。这里我们借鉴 ISO 规定（ISO 5492—1992）的典型食品质构的评价术语来加以介绍。

hardness（firmness）：硬度。表示使物体变形所需要的力。

cohesiveness：凝聚性。表示形成食品形态所需内部结合力的大小。

brittleness: 酥脆性。表示破碎产品所需要的力。

chewiness: 咀嚼性。表示把固态食品咀嚼成能够吞咽状态所需要的能量和硬度、凝聚性、弹性有关。

gumminess: 胶黏性。表示把半固态食品咀嚼成能够吞咽状态所需要的能量和硬度、凝聚性有关。

viscosity: 黏性。表示液态食品受外力作用流动时分子之间的阻力。

springiness: 弹性。表示物体在外力作用下发生形变，当撤去外力后恢复原来状态的能力。

adhesiveness: 黏附性。表示食品表面和其他物体（舌、牙、口腔）附着时，剥离它们所需要的力。

granularity: 粒状性。表示食品中粒子大小和形状。

Conformation: 组织性。表示食品中粒子的形状及方向。

moisture: 湿润性。表示食品中吸收或放出的水分。

fatness: 油脂性。表示食品中脂肪的量及质。

国际上定义的基本质构评价术语列于下面，供参考。

1、一般概念

structure: 结构、组织。表示物体或物体各组成部分关系的性质。

texture: 质构、质地。表示物质的物理性质（包括大小、形状、数量、力学、光学性质、结构）及由触觉、视觉、听觉的感觉性质。

2、与压缩、拉伸有关的术语

firm (hard): 硬。表示受力时对变形抵抗较大的性质（触觉）。

soft: 柔软。表示受力时对变形抵抗较小的性质（触觉）。

tough: 坚韧。表示对咀嚼引起的破坏有较强的和持续的抵抗性质。近似于质构术语中的凝聚性（触觉）。

tender: 柔韧。表示对咀嚼引起的破坏有较弱的抵抗性质（触觉）。

chewy: 筋道。表示像口香糖那样对咀嚼有较持续的抵抗性质（触觉）。

short: 脆。表示一咬即碎的性质（触觉）。

springy: 弹性。去掉作用力后变形恢复的性质（视觉）。

plastic: 可塑的。去掉作用力后变形保留的性质（视觉）。

sticky: 黏附性。表示咀嚼时对上颚、牙齿或舌头等接触面黏着的性质（触觉）。

glutinous: 黏稠状的。与发黏及黏附性视为同义语（触觉和视觉）。

brittle: 易破的。表示加作用力时，几乎没有初期变形而断裂、破碎或粉碎的性质（触觉和听觉）。

crumble: 易碎的。表示一用力便易成为小的不规则碎片的性质（触觉和视觉）。

crunchy: 咯蹦咯蹦的。表示兼有易破的和易碎的性质（触觉、视觉和听觉）。

crispy: 酥脆的。表示用力时伴随脆响而屈服或断裂的性质，常用来形容吃鲜苹果、芹菜、黄瓜、脆饼干时的感觉（触觉和听觉）。

thick: 发稠的。表示流动黏滞的性质（触觉和视觉）。

thin: 稀疏的。是发稠的反义词（触觉和视觉）。

3、与食品结构有关的术语

（1）颗粒的大小和形状。

smooth: 滑润的。表示组织中感觉不出颗粒存在的性质（触觉和视觉）。

fine: 细腻的。结构的粒子细小而均匀的样子（触觉和视觉）。

powdery: 粉状的。表示颗粒很小的粉末状或易碎成粉末的性质（触觉和视觉）。

gritty: 砂状的。表示小而硬颗粒存在的性质（触觉和视觉）。

coarse: 粗粒状的。表示较大、较粗颗粒存在的性质（触觉和视觉）。

lumpy: 多疙瘩状的。表示大而不规则粒子存在的性质（触觉和视觉）。

（2）结构的排列和形状。

flaky: 薄层片状的。容易剥落的层片状组织（触觉和视觉）。

fibrous: 纤维状的。表示可感到纤维样组织且纤维易分离的性质（触觉和视觉）。

strings: 多筋的。表示纤维较粗硬的性质（触觉和视觉）。

pulpy: 纸浆状的。表示柔软而有一定可塑性的湿纤维状结构（触觉和视觉）。

cellular: 细胞状的。主要指有较规则的空状组织（触觉和视觉）。

puffed: 膨松的。形容胀发得很喧腾的样子（触觉和视觉）。

crystalline: 结晶状的。形容像结晶样的群体组织（触觉和视觉）。

glassy: 玻璃状的。形容脆而透明固体状的。

gelatinous: 果冻状的。形容具有一定弹性的固体。觉察不出组织纹理结构的样子（触觉、视觉和听觉）。

foamed: 泡沫状的。主要形容许多小的气泡分散于液体或固体中的样子（触觉和视觉）。

spongy: 海绵状的。形容有弹性的蜂窝状结构样的（触觉和视觉）。

（4）与口感有关的术语

mouthfeel: 口感。表示口腔对食品质感感觉的总称。

body: 浓的。质构的一种口感表现。

dry: 干的。口腔游离液少的感觉。

moist: 潮湿的。口腔中游离液的感觉既不觉得少，又不感到多的样子。

wet: 润湿的。口腔中游离液有增加的感觉。

watery: 水汪汪的。因含水多而有稀薄、味淡的感觉。

juicy: 多汁的。咀嚼中口腔内的液体有不断增加的感觉。

oily: 油腻的。口腔中有易流动，但不易混合的液体存在的感觉。

greasy: 肥腻的。口腔中有黏稠而不易混合液体或脂膏样固体的感觉。

waxy: 蜡质的。口腔中有不易溶混的固体的感觉。

mealy: 粉质的。口腔中有干的物质和湿的物质混在一起的感觉。

slimy: 黏糊糊的。口腔中有黏稠而滑溜的感觉。

creamy: 奶油状的。口腔中有滑溜感。

astringent: 收敛性的。口腔中有黏膜收敛的感觉。

hot: 热的。有热的感觉。

cold: 冷的。有冷的感觉。

cooling: 清凉的。像吃薄荷那样，由于吸热而感到的凉爽。

第二节 标 度

在感官检验中，标度方法是感官体验的量化方式，通过这种数字化处理，感官评价可以成为基于统计分析、模型、预测等理论的定量科学。

标度方法广泛用于需要量化感觉、态度或喜好倾向性等各种场合，标度技术基于感觉强度的心理物理学模型。即增强物理刺激的能量或增加食品组分的浓度或含量，会导致其在感觉、视觉、嗅觉或味觉方面有多大程度的增强。

一、标度种类

有 4 种对事件的标度种类。通常是指名义标度、序级标度、等距标度和比率标度。这几

种标度是根据测量理论中测量水平提出的,适用于各个水平的各类统计分析和不同的建模水平。

1、名义标度

名义标度中,对于事件的赋值仅仅是作为标记。数值赋值仅仅是用于分析的一个标记、一个类项或种类,不反应序列特征。对这类数据的适当分析是进行频率计算并报告结果。对于不同产品或环境的不同反应频率,可通过卡方分析或其他非参数统计方法进行比较。利用这一标度各单项间的比较是说明它们是属于同一类别还是不同类别(相等与不相等的结果),而无法得到关于顺序、区别程度、比率或差别大小的结果。

2、序级标度

序级标度中,赋值是为了对产品的一些特性、品质或观点(如偏爱)标示排列的顺序。该方法赋给产品的数值增加表示感官体验的数量或强度的增加。例如,对葡萄酒的赋值可能根据感觉到的甜度而排序,或对香气的赋值可根据从喜爱到最不喜爱而排序。在这种情况下,数值并不说明关于产品间相对差别。排在第四的产品某种感官强度并不一定就是排在第一的产品的 $1/4$,它与排在第三的产品间的差别也不一定就和排在第三与第二的产品间的差别相同。所以,我们既不能对感知到差别的程度下结论,也不能对差别的比率或数量下结论。

排序的方法常见于感官偏爱研究中。许多数值标度法可能只产生序级数据,对此也存在很强的疑问,因为选项间的间距主观上并不是相等的。例如关于常用的市场研究标度,“极好——很好——好——一般——差”。这些形容词间的主观间距不均匀。评为好与很好的两个产品间的差别比评为一般和差的产品间的差别要小得多。但是,在分析时我们经常试图将1~5赋值给这些等级并取平均值,而且就像这些赋值数据反映相等的间距一样进行统计。一个合理的从极好到差的5点标度分析是计算各等级中反应者的数目,并进行频率比较。

通常,排序数据分析可以报告反应的中值作为主要趋势的概括,或者报告其他百分数以得到额外的信息,而包含加法和除法(例如平均数的计算)的数学运算并不恰当。

3、等距标度

当反应的主观间距相等时会出现等距标度。在该标度水平下,赋值的数据可以表示实际的差别程度。那么,这种差别度就是可以比较的,称为等距水平测量。用于感官科学的标度,几乎没有哪种能够满足有助于建立得到等距测量水平的检验,而且该水平经常又是假定的。明确支持这种水平的一种标度方法是用于喜爱——厌恶判断的9点类项标度,通常称为9点快感标度。这是一个平衡的标度法,所标明的反应选项有大致相等的间距,如下所示:

非常喜欢
很喜欢
一般喜欢
有些喜欢 ,
既不喜欢也不厌恶
有些厌恶
一般厌恶
很厌恶
非常厌恶

这些选项通常从1~9,或者等间距(例如-4~+4)赋值进行编码和分析。对各种形容词标识的标度点间明确间距的大量研究导致了选择这些选项。可以看到,上述标度的区别在于强调主体反应的形容词,这些词汇会来自英语国家。在其他国家是否表现良好的等距,有待进一步试验。

4、比率标度

比率标度，在这种方式下，0 点不是任意的，而且数值反应了相对比例。

二、常用标度方法

有三种常用的标度方法。最古老也是最广为使用的标度方法也许是类项评估,评价员根据特定而有限的反应,将数值赋予觉察到的感官刺激。第二种方法与此相对应,是量值估计法,这种方法评价员可以对感觉赋予任何数值来反映其比率。第三种常用方法是线性标度法。该方法是评价员采用在一条线上作标记来评价感觉强度或喜爱程度。

这些方法存在两大方面的差异，首先是评价员所允许的自由度及对反应的限制。开放式标度法不设上限，其优点是允许评价员选择任何合适的数值进行标度。不过，这种开放式的反应难以在不同评价员之间进行校准。数据编码、分析及翻译过程会复杂化。相反，简单的分类法则易于确定固定值或使参照标准化，便于校准评价员，而且数据编码与分析常常很直观。标度方法间差异的第二个方面是允许评价员的区别程度。有的允许评价员根据需要任意使用任意多个中间值，而有的则被限制只能使用有限的离散的选择。所幸的是，采用合适的标度点数量似乎可以减少这些差异。9点（或更多）类项标度法与分级更精细的量值估计法及线性标记法结果很接近，尤其是当产品差异不是很大时。

1、类项标度

类项标度与线性标度的差别在于人们的选择受到很大的限制。图表标度技术至少给人的印象是——反应是连续分级的。实际上,作为数据,由于数据编码器具,它们同样也限制在不连续的可测量选项上,例如利用数字转换器的处理或通过光笔在触摸屏 CRT 上的像素分解的数值。但是在类项标度中,可选择的反应数目通常要少得多,典型的为 7~15 个类项。类项的多少取决于实际需要以及评价员对产品能够区别出来的级别数。随着评价人员训练的进展,对强度水平可感知差别的分辨能力会得到提高。

类项标度有时也被通称为“评估标度”，尽管这个术语也用于指所有的标度方法。最简单的，也是历史上最常见的形式是利用整数来反映逐渐增强的感官强度。作为一种心理物理学方法，它在早期文献中以“单刺激方法”出现，但与其他比较技术相比，它较少被用于测量绝对和辨别阈。不过，直接标度被认为是最经济的。也就是说，尽管他们可能是“廉价的数据”，但他们的优势在于对一个单一刺激，你至少可以得到一个数据点。如果是剖面或描述性分析，你可以得到许多数据点，标度了各个特征。在实际的感官工作中，由于检验完成时间的快慢和有竞争力的产品开发的紧迫性通常是一个重要因素，因此，对这种经济特性似乎要超过对标度有效性的关注。简单类项标度的例子如图所示。

类项标度举例:

(a) 数值 (整数) 标度 (根据 Lawless 和 Malone, 1986);

例	强度	1	2	3	4	5	6	7	8	9
		弱								强

(b) 语言类项标度 (根据 Macredy 等人, 1974):

例	氧化痕量 极微量 微量 少量 中等	无感觉 不确定
---	-------------------------------	------------

[illegible]

(e) 整体差异类项标度 (根据 Aust 等人, 1985):

例 与参照的差别

 无差别

 差别极小

 差别很小

 差别中等

 差别较大

 差别极大









太好了 很好 好 可能好或不好 差 很差 太差了

这类标度方法应用广泛。常见的例子是用大约 9 个点的整数反应。分级也可以更多，如 Winakor 等人用 1~99 的选项来评估织物的手感特征。在频谱方法中研究者使用了 15 点类项，但允许把每类再分成 10 等份，理论上相当于变成了至少 150 点的标度。在快感或情感检验中，常用两极标度，有一个 0 点或中性点位于中间位置。这比强度标度简略，例如，在对儿童使用的“笑脸”标度中，虽然对较年长的儿童可以使用 9 点法，但对很年幼的受试对象只采用 3 个选择。在以后的研究中，有研究者放弃使用标度或整数，以避免受试对象的偏见，因为人们常对特定的数字产生特定的含意。为解决这一问题，采用未标注的方格标度法。

数字类项标度中可用数字有上限的事实便于线性等距标度的实现。如果应用合理,类项

标度可以很接近等距测量。一个主要问题是要提供足够的选择，以表示评价人员能够分辨的差别。

如果评价小组受过高度训练而能够辨别很多刺激水平的话，一个简单的 3 点标度是不够的。在风味剖面标度中，可能逐渐使用很多的标度点。然而，有规律表明使用太多的标度点会降低回收率；进一步的详细研究可以将产品更好地区别到某一点上，而额外的反应选择只不过是记录了随机误差波动，从而使回收率降低。

有一个个人倾向问题，尤其在消费者工作中，会以去除选项或截去端点的方法来简化标度。回避端点会引起不良结果。一些受试者往往不愿使用端点类项，以防在后面的检验中出现更强或更弱的情况，因此，人们自然倾向于避免端点类项，而将 9 点标度截为 7 点标度，可能使评价者实际上只能得到 5 点标度的作用。所以，最好避免在实验计划中截去标度点的倾向。

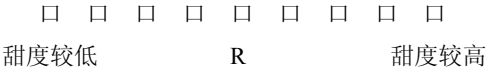
实验设计者要考虑是否对中间标度点给予物理学示例。给出事例在端点类项中常见，而在中间点类项中的使用要少一些。这种做法的优点是可以获得一定的标准水平，这是受过训练的有描述能力的评价人员所希望的特征。因此，我们发现在质地剖面方法中用 9 点标度评价硬度时，所有的类项点都有相应的物理学示例，从非常软的像白色的煮鸡蛋那样到非常硬的像硬糖块那样，两个相邻的样品间保持的间距大致相等。

事例潜在的缺点是限制了受试者的标度。在实验者看来是等距的，在参与者看来却未必。这种情况下，最好的做法是让受试者在可选择标度范围内确定判断，但不假设中间点的示例是真正等距的。这一选择取决于实验者，其决定反映了他是更关心希望得到标准还是更关心反映限制或潜在偏差。训练有素的评价小组进行感官分析时更喜欢标准，而消费者偏爱评估时则喜欢减少限制。

在实践中，简单的类项标度对产品区别的敏感性几乎和其他标度技术，包括线性标度法和量值估计法一样。由于它们的简易性，所以特别适合于消费者工作。另外，它们在快速准确的数据编码和列表方面也有一些优势，因为它们的工作量要小于线性标度或变化更多的可能包含分数的量值估计法。当然，前提假设数据列表是手工进行的，如数据是利用计算机系统在线记录的，就不存在这一优势。具有固定类项的多种标度现在仍在使用，包括用于观点和态度的利开特(Liken)式标度，它的类项是基于人们对关于该产品的表述同意与否的程度，类别选项对很多情况具有灵活性。

下面例子是表示类项标度数据的 t 检验分析。

例：某巧克力奶生产商希望比较一下它的产品（“自有”品牌）与当地一家竞争者和一家全国性品牌的奶粉产品的甜度水平。利用 9 点标度，各种乳品与其自有品牌相比较，标示为 R 的中点作为参照表示如下：



16 名巧克力奶的消费者参与了检验。数据编码为从 1~9 分，t 检验相对于数值 5（参考点）进行处理。参照也相对于自身（盲标）检验该方法的准确性。数据及统计结果列于下表：

评审员	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	总计	均值	平方和	标准方差	标准误差
全国样 X	6	9	6	3	6	6	8	7	7	5	4	6	8	4	3	8	96	6.00	626	1.826	0.456
当地样 Y	3	7	5	6	4	4	5	3	5	5	6	4	4	5	2	6	74	4.63	368	1.310	0.328
自有样 Z	6	5	4	6	5	5	5	5	5	4	5	5	5	4	6	7	82	5.13	130	0.806	0.202

则：

$$t = (\text{平均值} - 5.0) / \text{标准误差}$$

$$t = (6.00 - 5.0) / 0.456 \quad (4.63 - 5.0) / 0.328 \quad (5.13 - 5.0) / 0.202$$

该试验的无差异假设是人群样本对产品甜度的平均值 5.0。对立假设是不等于 5.0（可能高于或低于，该检验是双侧的）。

16 名小组评价员自由度为 15，所以 α 风险为 0.05 时双侧检验的临界值是 2.131。如果 $t > 2.131$ ，当无差异假设成立时，可以期望存在较大的差别或远低于 5% 的次数。

结果如下：

全国性竞争者	当地竞争者	“自由”竞争者
否定无差异	否定失败	否定失败

结果表明全国性竞争者的产品较我们自有品牌的甜，而要得出关于当地竞争者产品的结论却没有充分的证据。我们可以试着下结论，我们的自有品牌和当地竞争者的产品甜度是相似的。

2、量值估计

量值估计方法是流行的标度技术，它不受限制地应用数字来表示感觉的比率。在此过程中，评价员允许使用任意正数并按指令给感觉定值，数字间的比率反映了感觉强度大小的比率。例如，假设产品 A 的甜度值为 20，产品 B 的甜度是它的 2 倍，那么 B 的甜度评估值就是 40。量值估计不依赖于像类项标度和线性标度技术中那样可见到选票。

应用这种方法需要注意对受试者的指令以及数据分析技术。

量值估计有两种基本变化形式。第一种形式，给受试者一个标准刺激作为参照或基准，此标准刺激一般给它一个固定数值。所有其他刺激与此标准刺激相比较而得到标示，这种标准刺激有时称为“模数”。另一种主要的变化形式则不给出标准刺激，参与者可选择任意数字赋予第一个样品，然后所有样品与第一个样品的强度比较而得到标示。实践中受试者可能“一环套一环”地根据系列中最靠近的一个样品给出评估。

在心理物理学实验室，量值估计得到了初步应用，一般每次只标度一种属性。但是，评估多个特征或剖面分析也被用于味觉研究，这种方法也自然地被沿用到具有多重味道和芳香特征的食品研究中。

参照样或赋以固定数值模数的量值估计应用示范指令如下：

请品尝第一个样品并注意其甜度。这是一个参照样品，它的甜度值定为“10”。请根据该参照样品来评价所有其他样品，给这些样品相应的数值以表示样品间的甜度比率。例如，如果下一个样品的甜味是参照样的 2 倍，则将其定值为“20”，如果其甜度是参照样的一半，则将其定值为“5”，如果其甜度是 3.5 倍，则将其定值为“35”。你可以使用任意的正数，包括分数和小数。

在这种方法中有时允许用数字 0，因为在检验时有些产品实际上没有甜味，或者没有需评价的感官特性。但参照样品不能用 0 来赋值，参照样最好能选择在强度范围的中间点附近。没有感觉特征的产品定值为 0 可以理解，但使数据分析复杂化了。

量值估计的另一个主要变化形式不使用参考点。这种情况指令如下：

请品尝第一个样品并注意其甜度。请根据该参照样来评价所有其他样品，并给这些样品相应的数字以表示样品间甜度的比率。例如，如果下一个样品的甜味是参照样的 2 倍，则给该样品定值为第一个样品的 2 倍；如果甜味是参照样的一半，则给其定值为第一个样品的一半；如果甜味是 3.5 倍，则给其定值为 3.5 倍。你可以使用任意正数，包括分数和小数。

参与者一般会选择他们感觉合适的数字范围，ASTM 法建议第一个样品的值在 30~100 之间为宜，应避免使用太小的数字。参与者应注意避免前面使用有界限类项标度的习惯，如

限制数字范围为 0~10。这对于以前受过训练使用其他标度方法的评价人员是一个很大的困难，因为他们总是习惯于坚持了解的而感到习惯的方法。有这种行为的评价人员可能没有理解指令中“比率”的特性。为避免这一问题，可以让参与者进行一些准备活动来帮助他们确切地理解标度指令。准备活动可以让他们估计不同几何图形的大小和面积或者线段的长度。有时，要求评价人员同时标度多个特征或将整体强度分解为特定的属性。如果需要这种“剖面”，几何图形可以包含不同的阴影区域，或者不同颜色的线段。

如果允许参加者选择自己的数字范围，那么，在统计分析之前有必要进行再标度，使每个人的数据落在一个正常的范围内。这样，可以防止受试对象选择极大数字而对集中趋势(平均值)测量和统计检验的不良影响。这一再标度过程也被称为“标准化”(ASTM, 1995)。再标度的一种常用方法是：①计算每个人全部数据的几何平均值；②计算所有数据(将全部受试者综合起来)的总几何平均值；③对各受试者计算总几何平均值与各自几何平均值的比率，由此得到各受试者的再标度因子，构建这一因子也可以不用总几何平均值，而选用任何正数，例如选用数值 100；④对于各受试者，用他们各自的数据点乘以他们相应的再标度因子。这样，产品就可以进行统计学比较，并得到集中趋势量度。如果数据在再标度前已经转化成对数值，那么，再标度因子则是基于对数的平均值，这样，它就变为用加法而不是用乘法。

量值估计的数据常常在数据分析前转换成对数，这主要是因为数据趋向于对数常态分布，或者至少是正偏离。在标度中有一些高度偏离值，而大部分标度位于较低的数值范围内。原因是标度在顶端是开放的，而在底部则以零为界。不过，当数据中包含零的时候，将数据转换成对数和几何平均值也会出现一些问题。0 的对数是没有意义的，而在用乘法计算 N 次几何平均值时也将使结果为 0。对于这个问题有几种办法。一种方法是将数据中的 0 赋予一个小的正数，比如取受试者给出的最小标度值的一半 (ASTM, 1995)。当然，结果分析会受这种选择的影响。另一种方法是在计算标准化因子时使用算术平均值或中间值。对于再标度它是可行的，但并未去除数据的偏离。

在实践中，量值估计法可应用于训练有素的评价小组、消费者甚至是儿童。但是，比起受到限制的标度方法，量值估计法的数据变化更大，特别是出自未经训练的消费者之手的数据。该标度法的无界限特性，使得它特别适合于那些上限会限制评价人员在评估感官特征中区分感官体验的能力的情况。例如，像辣椒的辣度这样的刺激或痛觉，在类项标度法中可能都被评估为接近上限的强度。但在端点开放的量值估计法中，允许评价人员有更大的自由度来运用数字反映极强烈的感觉变化。

在喜爱和厌恶的快感标度中，使用量值标度还要考虑一个问题。这种技术的应用有两种选择，一种使用单侧或单极标度来表示喜爱的程度，另一种使用双极标度，可以使用正数和负数，外加一个中性点。在喜爱和厌恶的双极量值标度中，允许使用正数和负数来表示喜爱和厌恶的比率或比例。对正数和负数的选择只表示数字代表的是喜欢还是不喜欢。在单极量值估计中，则只允许使用正数(有时包括 0)。低端表示厌恶，随着数值的增大，表明喜爱的程度成比例逐渐升高。设计这种标度时，实验者应明确单极标度对参与评价的人员是否合适，因为没有认识到事实上存在中性反应的情况，也没有认识到存在明显的两种反应方式，即喜欢和不喜欢。如果能保证所有的结果都在快感的一侧——无论是都喜欢或都不喜欢，只是程度不同，那么单极标度才有意义。在少数情况下，对食品或消费产品的检验是可以采用的。这时，至少某些参与者可忽略变化或者观点的改变不明显。因此，像 9 点类项标度的双极标度更符合常识。

3、线性标度

线性标度也称为图表评估标度或视觉相似标度。自从发明了数字化设备以及随着在线计

计算机化数据输入程序的广泛应用，这种标度方法变得特别流行。其基本思想是让评价员在一条线段上做标记以表示感官特性的强度或数量。大部分情况下，只有端点做了标示。标示点也可以从线段两端缩进一点儿以避免末端效应。其他中间点也可以标出来。一种常见的变化形式是标出一个中间的参考点，代表标准品或基线产品的标度值。所需检验的产品根据此参考点来进行标度，如图所示，经过训练的评价员对多特性进行描述性分析时，这些技术是很常用的。而在消费者研究中则较少应用。

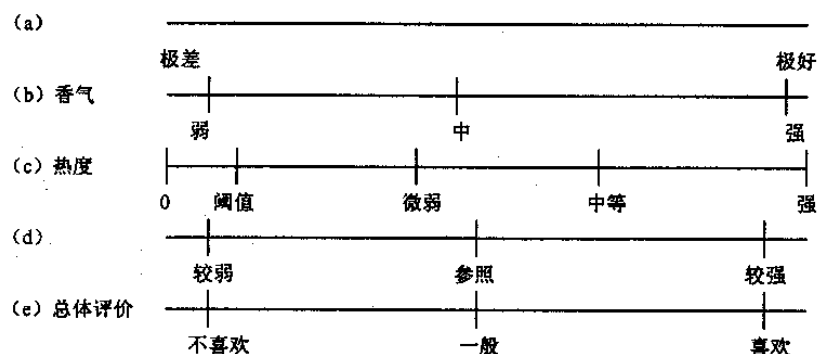


图 5-3 线性标度

(a) 端点标示 (根据 Baten, 1946); (b) 端点缩进 (根据 Mecredy, 1974);

(c) ASTM 法 1083 中的附加点标示 (ASTM, 1991); (d) 利用直线的相对参考点标度;

(e) 利用直线的快感标度 (根据 Stone 等人, 1974)

感官评价的线性标度起源于二次大战中美国密歇根州农业实验站的一次实验。在 Baten 的研究中，检验了苹果的各种贮存温度对水果的吸引力进行了简单的类项标度（从很理想到很不理想分成 7 个选项进行评估），又使用了 6 英寸的线性标度，线的左端标示为“极差”，线的右端标示为“极好”。对显示在线上的反应用英寸为单位进行测量。

Baten 的研究说明了以前的文献中许多研究者的观察敏锐性有多大。线性标度提供了选项的连续等级选择，只受限于数据列表的测量能力。

Stone 等人 (1974) 推荐将线性标度用于定量描述分析 (Quantitative Descriptive Analysis, QDA)，继而形成一种标示各种重要感官特性的新方法。在定量描述分析 (QDA) 中，使用一种近似于等距标度的标度方法是很重要的，因为在描述分析中方差分析已成为比较产品的标准统计技术。

自 QDA 出现以来，线性标度技术已被用于需要感官反应的各种不同场合。例如，Einstein (1976) 成功地使用线性标度法让消费者评价啤酒的风味强度、丰满度、苦味和后味特征。所谓“成功”是指在检验实例中获得了统计上的显著差别。线性标度法的应用并不局限于食品和消费产品，在临床上对痛觉和祛痛的度量也可利用水平线或垂直线进行线性标度法，Lawless (1977) 用线性标度技术并结合比率指令对口味和气味的强度及快感判断进行了混合研究。这是一个综合的方法，受试者根据指令进行线性标度，如同进行量值评估一样。比如说，如果一种产品的甜度是前一种的 2 倍，那么就在线上 2 倍距离的位置标示出来。这种情况下，线的端点是一个问题，受试者被告知如果空间不够的话可以另附纸，但很少有人会这样做。在比较类项标度、线性标度和量值评估时，线性标度方法对产品的差别与其他标度技术几乎同样灵敏。